View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

higher than that of the first visit ($p < 0.000$). We tested the null that the coefficient on the number of visit is equal to zero. We rejected the null hypothesis that the coefficient on the number of visits dummy variable is equal to zero ($p < 0.000$). **CONCLUSIONS:** Our study finds that return visits are associated with higher patient satisfaction than first visit. Furthermore, the dummy variable of the number of visit explains patient satisfaction score well.

PMC43

EIGHT KEY EVIDENCES FOR CONTENT VALIDITY OF PATIENT REPORTED OUTCOME (PRO) INSTRUMENTS

Martin ML
Health Research Associates, Inc, Seattle, WA, USA

OBJECTIVES: The FDA Guidance on the development and use of Patient Reported Outcome (PRO) instruments emphasizes the need to establish content validity for the concepts addressed by the PRO items. Content validity is the relationship between an instrument's content and the construct it intends to measure. This paper presents eight key points of criteria for supporting evidences of content validity for PRO instruments. **METHODS:** Language is the primary means of patient recognition and expression of concepts. Concept relevance is assessed by qualitative analysis. A Concept Elicitation process is used to describe the relationship between concepts and patient experience. Patient comprehension of a concept is evaluated through cognitive interviews. The design of interviews and treatment of results must focus on specific criteria in order to support content validity of the PRO. **RESULTS:** The eight key criteria required to support PRO content validity are: (1) concepts are relevant to the patient experience with their condition; (2) qualitative data collection has been conducted to the point of "saturation of concept" (3) concepts are presented in the langua ge of the patients; (4) the appropriate aspect of the concept is being evaluated; (5) the presentation of the concept in a PRO can be properly comprehended by patients; (6) response options are meaningful and clear; (7) the recall period is appropriate to the patient experience of their condition; and (8) the concepts and language used in the PRO are adaptable for use in global trials. This paper presents examples of data collection, and documentation for each criterion. **CONCLUSIONS:** Evidences of content validity are can only be provided by qualitative data. The qualitative interview process may vary in specific methods but should adhere to consistent objectives, and should cover at least eight key evidences in order to provide adequate support for content validity of a PRO.

PMC44

AN EVALUATION OF THE STATISTICAL EFFICIENCY OF THE US POPULATION MEDIAN-BASED EQ-5D INDEX USING DATA FROM THE MEDICAL EXPENDITURE PANEL SURVEY

Luo X, Shaw JW, Pickard AS, Walton S
University of Illinois at Chicago, Chicago, IL, USA

OBJECTIVES: An algorithm was recently developed to predict median US population preferences for EQ-5D health states. The primary objective of this study was to provide evidence regarding the estimation efficiency of the median-based EQ-5D index relative to the existing mean-based index. A secondary objective was to evaluate the sensitivity of findings to the application of robust statistical procedures. **METHODS**: Data were taken from the 2002–2003 Medical Expenditure Panel Survey. Adult survey participants completed the EQ-5D and 12-item Short-Form Health Survey (SF-12); rated their physical and mental health; and reported the presence of chronic conditions. Associations of mean- and median-based EQ-5D index scores with health status measures and chronic conditions were estimated using least squares (LS) and rank (R) regression. Associations of changes in index scores with changes in health status measures were similarly estimated. For each set of analyses, a relative efficiency (RE) statistic was derived as the ratio of model F statistics (median-based index/mean-based index). **RESULTS**: RE statistics (LS/R) for associations of index scores with health status measures and chronic conditions were as follows: SF-12 mental summary, 0.50/1.35; SF-12 physical summary, 0.40/0.75; mental health rating, 0.63/1.10; physical health rating, 0.54/0.99; angina, 0.55/0.89; arthritis, 0.50/0.88; asthma, 0.77/0.98; coronary heart disease, 0.54/0.91; diabetes, 0.60/0.95; emphysema, 0.65/1.03; hearing problems, 0.50/0.98; myocardial infarction, 0.51/0.97; stroke, 0.66/1.00; and visual problems, 0.65/1.03. RE statistics (LS/R) for associations of index score changes with changes in SF-12 physical and mental summaries were 0.63/0.89 and 0.59/1.17, respectively. **CONCLUSIONS**: Parametric analyses exhibited reduced power when applied to the median-based index scores. However, robust procedures suffered from minimal efficiency losses and exhibited efficiency gains when used to analyze associations of index scores with mental health measures. These findings reflect the difference in skewness between the mean- and median-based index scores, which can be attributed to the difference in score ranges.

CONCEPTUAL PAPERS & RESEARCH ON METHODS – Statistical Methods

PMC45

INTEGRATED INTERPRETATION OF KAPPA WITH SENSITIVITY AND SPECIFICITY

Shimbo T[1], Miyaki K[1], Sakai M[1], Takahashi O[2], Ishizuka N[1]
[1]International Medical Center of Japan, Tokyo, Japan, [2]St.Luke's Life Science Institute, Tokyo, Japan

OBJECTIVES: The reproducibility (kappa) and accuracy (sensitivity/specificity: SN/SP) of diagnostic tests has often been reported. However, both values are discussed separately, even when simultaneously determined. Because good reproducibility is the premise for good accuracy, integrated interpretation of both is required. The present study investigates the relationship between kappa and measured SN/SP by Monte-Carlo simulation given true accuracy and measurement error. METHODS: We assumed diseased and non-diseased populations with test values that are normally distributed with different means and the same standard deviation. True SN/SP given by a threshold value is decreased by the measurement error that is also normally distributed and this leads to measured SN/SP. We assumed that sensitivity equals specificity for convenience. Two measurements can generate kappa statistics. We simulated this process using Monte-Carlo simulation. The true mean difference of both the population and the standard deviation of measurement error was changed for a variety of combinations then kappa and measured SN/SP were compared. We assumed a total population of 100,000, and disease prevalences of 10% and 50%. Simulation was done in STATA 10/SE. RESULTS: Disease prevalence did not substantially affect the results, so only the results for 10% prevalence are described. At a true SN/SP of 93.1% and kappa values of 0.75, 0.5 and 0.25, measured SN/SP decreased to 91.3%, 86.1% and 77.1%, respectively. At a true SN/SP of 68.9% and kappa values of 0.75, 0.5, 0.25, measured SN/SP decreased to 68.3%, 66.0%, 61.7%, respectively. Now we know measured SN/SP can substantially or not improve so much when kappa improves, depending on the combination of kappa and measured SN/SP. CONCLUSIONS: The relationship between kappa and SN/SP allows the integrated interpretation of both, and indicates how much accuracy can be improved by improving reproducibility.

PMC46

COMPARISON OF DIFFERENCE-IN-DIFFERENCE AND PROPENSITY SCORE METHODS IN REGRESSIONS WITH BINARY OUTCOMES

Cao Z[1], Margolis J[2], Mark T[3]
[1]Thomson Reuters, Cambridge, MA, USA, [2]Thomson Reuters, Bala Cynwyd, PA, USA, [3]Thomson Reuters, Washington, DC, USA

OBJECTIVES: Difference-in-difference (DID) and propensity score matching (PSM) methods are two ways of evaluating the effect of changes in policies influencing drug prescribing or medications. The objective of this study is to compare and contrast the strengths and weakness and different types of information provided by the DID and PSM methods in the context of determining the impact of an insurance plan restriction on binary health care utilization outcomes. Adult patients diagnosed with diagnosis A were selected from MarketScan® Commercial Database (2005–2007), a large claims data form persons with employ-based insurance. **METHODS:** The DID analysis was implemented by a GEE regression with logit link, binomial variance function, and exchangeable correlation structure on panel data with pre- and post-restriction period. The dependent variables included inpatient hospitalization, ER visit, use of alternative drug B, and use of alternative drug C. The key independent variables included an indicator for plan restriction, an indicator for pre- or post- restriction period, and an interaction term between them. Demographic and clinical characteristics were included as covariates. The odds ratios (OR) of the interaction terms were considered as the DID estimates. In the PSM method, patients in restricted and unrestricted plans were propensity score matched based on pre-implementation characteristics, followed by logistic regressions. **RESULTS:** Without controlling for any covariate, the restriction effects were non-significant on all outcomes using DID analysis. The restriction effects were significant for ER visit and hospitalization using PSM. When all characteristics were controlled, the restriction effects were non-significant in both methods, except the outcome of drug B use. **CONCLUSIONS:** DID and PSM- based methods may produce different estimates. Neither method can fully control all potential confounders. Compared with the PSM method, DID method uses same individual as the control, but is restricted to studies when pre- and post-event data are available.

PMC47

STATISTICAL METHODS NEEDED IN HEALTH ECONOMICS

Gause D, Gause P
See-It-Analytics, Easton , PA, USA

OBJECTIVES: To review important statistical methods used in health economics. **BACKGROUND:** During two decades in health economics at Novartis Pharmaceuticals, statistician Doug Gause regularly presented friendly one-page illustrations of statistical concepts in health economics. Thirty of these "Analysis Notes" are presented on this poster, along with a magnifying glass for easier viewing. **ANALYSIS NOTES:** *Making inferences*: Confidence interval estimation instead of hypothesis testing, and bootstrapping to avoid assumptions. A*nalyzing costs*: Two-part GLM models instead of transformations. *Measuring compliance*: Avoiding adherence bias in calculating MPR. Competing risks of discontinuing, adding, switching, or remaining on drug. *Data visualization:* Regression trees, drug-o-grams, mosaic plots, parallel coordinate plots, and sequence plots to visualize patterns. (Lets reserve pie charts for the Baker's Convention !) *Covariate adjustment*: Marginal means/IPTW, confounders vs. colliders, instrument variables, and time dependent covariates to incorporate changes over time. *Historical data:* Bayesian analysis for bias adjusting, and meta-analysis to incorporate results from multiple studies. *Avoiding pitfalls*: Be aware of the "individualized" fallacy and the "ecological" fallacy. *Differential follow-up*: Person-time and survival analysis to accommodate censoring. *Predictive modeling:* Nearest neighbor vs. recursive partitioning vs. neural networks. *Project planning*: Sample size to handling multiplicity. *Clustered/longitudinal data:* Random effects models to account for correlated observations from the same physician or repeated measurements from the same patient. *Cost-effectiveness:* Angles instead of ratios, and incremental net benefit to handle multiple effectiveness measures. *Heterogeneous data:* Quantile and polytomous regression to compare different MPR distributions. **CONCLUSIONS:** Statistics